
Sequence Listing was accepted.

If you need help call the Patent Electronic Business Center at (866) 217-9197 (toll free).

Reviewer: Anne Corrigan

Timestamp: [year=2007; month=12; day=14; hr=9; min=59; sec=44; ms=179;]

Validated By CRFValidator v 1.0.3

Application No: 10031496 Version No: 7.0

Input Set:

Output Set:

Started: 2007-11-26 18:43:03.925

Finished: 2007-11-26 18:43:08.778

Elapsed: 0 hr(s) 0 min(s) 4 sec(s) 853 ms

Total Warnings: 99
Total Errors: 0

No. of SeqIDs Defined: 99

Actual SeqID Count: 99

| Error code | | Error Description |
|------------|-----|---|
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (1) |
| W | 402 | Undefined organism found in <213> in SEQ ID (2) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (3) |
| W | 402 | Undefined organism found in <213> in SEQ ID (4) |
| W | 402 | Undefined organism found in <213> in SEQ ID (5) |
| W | 402 | Undefined organism found in <213> in SEQ ID (6) |
| W | 402 | Undefined organism found in <213> in SEQ ID (7) |
| W | 402 | Undefined organism found in <213> in SEQ ID (8) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (9) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (10) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (11) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (12) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (13) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (14) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (15) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (16) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (17) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (18) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (19) |
| W | 213 | Artificial or Unknown found in <213> in SEQ ID (20) |

Input Set:

Output Set:

Started: 2007-11-26 18:43:03.925 **Finished:** 2007-11-26 18:43:08.778

Elapsed: 0 hr(s) 0 min(s) 4 sec(s) 853 ms

Total Warnings: 99

Total Errors: 0

No. of SeqIDs Defined: 99

Actual SeqID Count: 99

| Error code | Error Description |
|------------|--|
| W 213 | Artificial or Unknown found in <213> in SEQ ID (21) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (22) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (23) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (24) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (25) |
| W 213 | Artificial or Unknown found in <213> in SEQ ID (26) This error has occured more than 20 times, will not be displayed |
| W 402 | Undefined organism found in <213> in SEQ ID (98) |
| W 402 | Undefined organism found in <213> in SEQ ID (99) |

SEQUENCE LISTING

```
<110> National Renewable Energy Laboratory (NREL)
<120> Cellobiohydrolase I Gene and Improved Variants
<130> NREL 99-45
<140> 10031496
<141> 2002-01-14
<160> 99
<170> PatentIn version 3.4
<210> 1
<211> 45
<212> DNA
<213> Artificial
<220>
<223> Nucleotide encoding linker
<400> 1
cctcccggcg gaaacccgcc tggcaccacc accacccgcc gccca
                                                                    45
<210> 2
<211> 15
<212> PRT
<213> Trichoderma reesei
<400> 2
Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr Thr Arg Arg Pro
                                   10
                                                      15
<210> 3
<211> 24
<212> DNA
<213> Artificial sequence
<220>
<223> Nucleotide encoding linker
<400> 3
                                                                    24
ggcggaaacc cgcctggcac cacc
<210> 4
<211> 1551
<212> DNA
<213> Trichoderma reesei
```

<400> 4

| atgtatcgga | agttggccgt | catctcggcc | ttcttggcca | cagctcgtgc | tcagtcggcc | 60 |
|------------|------------|------------|------------|------------|------------|------|
| tgcactctcc | aatcggagac | tcacccgcct | ctgacatggc | agaaatgctc | gtctggtggc | 120 |
| acgtgcactc | aacagacagg | ctccgtggtc | atcgacgcca | actggcgctg | gactcacgct | 180 |
| acgaacagca | gcacgaactg | ctacgatggc | aacacttgga | gctcgaccct | atgtcctgac | 240 |
| aacgagacct | gcgcgaagaa | ctgctgtctg | gacggtgccg | cctacgcgtc | cacgtacgga | 300 |
| gttaccacga | gcggtaacag | cctctccatt | ggctttgtca | cccagtctgc | gcagaagaac | 360 |
| gttggcgctc | gcctttacct | tatggcgagc | gacacgacct | accaggaatt | caccctgctt | 420 |
| ggcaacgagt | tctctttcga | tgttgatgtt | tcgcagctgc | cgtgcggctt | gaacggagct | 480 |
| ctctacttcg | tgtccatgga | cgcggatggt | ggcgtgagca | agtatcccac | caacaccgct | 540 |
| ggcgccaagt | acggcacggg | gtactgtgac | agccagtgtc | cccgcgatct | gaagttcatc | 600 |
| aatggccagg | ccaacgttga | gggctgggag | ccgtcatcca | acaacgcgaa | cacgggcatt | 660 |
| ggaggacacg | gaagctgctg | ctctgagatg | gatatctggg | aggccaactc | catctccgag | 720 |
| gctcttaccc | cccacccttg | cacgactgtc | ggccaggaga | tctgcgaggg | tgatgggtgc | 780 |
| ggcggaactt | actccgataa | cagatatggc | ggcacttgcg | atcccgatgg | ctgcgactgg | 840 |
| aacccatacc | gcctgggcaa | caccagcttc | tacggccctg | gctcaagctt | taccctcgat | 900 |
| accaccaaga | aattgaccgt | tgtcacccag | ttcgagacgt | cgggtgccat | caaccgatac | 960 |
| tatgtccaga | atggcgtcac | tttccagcag | cccaacgccg | agcttggtag | ttactctggc | 1020 |
| aacgagctca | acgatgatta | ctgcacagct | gaggaggcag | aattcggcgg | atcctctttc | 1080 |
| tcagacaagg | gcggcctgac | tcagttcaag | aaggctacct | ctggcggcat | ggttctggtc | 1140 |
| atgagtctgt | gggatgatta | ctacgccaac | atgctgtggc | tggactccac | ctacccgaca | 1200 |
| aacgagacct | cctccacacc | cggtgccgtg | cgcggaagct | gctccaccag | ctccggtgtc | 1260 |
| cctgctcagg | tcgaatctca | gtctcccaac | gccaaggtca | ccttctccaa | catcaagttc | 1320 |
| ggacccattg | gcagcaccgg | caaccctagc | ggcggcaacc | ctcccggcgg | aaacccgcct | 1380 |
| ggcaccacca | ccacccgccg | cccagccact | accactggaa | gctctcccgg | acctacccag | 1440 |
| tctcactacg | gccagtgcgg | cggtattggc | tacagcggcc | ccacggtctg | cgccagcggc | 1500 |
| acaacttgcc | aggtcctgaa | cccttactac | tctcagtgcc | tgtaaagctc | С | 1551 |

<210> 5

<211> 514

<212> PRT

<213> Trichoderma reesei

Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg 1 5 10 15

Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr 20 25 30

Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser 35 40 45

Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser 50 55 60

Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp 65 70 75 80

Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala 85 90 95

Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe $100 \,$ $105 \,$ $110 \,$

Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met 115 120 125

Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe 130 135 140

Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro \$165\$ \$170\$ \$175\$

Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln
180 185 190

Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly 195 200 205

Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly
210 215 220

| Ser 225 | Суз | Суз | Ser | Glu | Met 230 | Asp | Ile | Trp | Glu | Ala 235 | Asn | Ser | Ile | Ser | Glu 240 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ala | Leu | Thr | Pro | His 245 | Pro | Суз | Thr | Thr | Val 250 | Gly | Gln | Glu | Ile | Cys 255 | Glu |
| Gly | Asp | Gly | Cys 260 | Gly | Gly | Thr | Tyr | Ser 265 | Asp | Asn | Arg | Tyr | Gly 270 | Gly | Thr |
| Cys | Asp | Pro 275 | Asp | Gly | Суз | Asp | Trp 280 | Asn | Pro | Tyr | Arg | Leu 285 | Gly | Asn | Thr |
| Ser | Phe 290 | Tyr | Gly | Pro | Gly | Ser 295 | Ser | Phe | Thr | Leu | Asp 300 | Thr | Thr | Lys | Lys |
| Leu 305 | Thr | Val | Val | Thr | Gln 310 | Phe | Glu | Thr | Ser | Gly 315 | Ala | Ile | Asn | Arg | Tyr 320 |
| Tyr | Val | Gln | Asn | Gly 325 | Val | Thr | Phe | Gln | Gln 330 | Pro | Asn | Ala | Glu | Leu 335 | Gly |
| Ser | Tyr | Ser | Gly 340 | Asn | Glu | Leu | Asn | Asp 345 | Asp | Tyr | Cys | Thr | Ala 350 | Glu | Glu |
| Ala | Glu | Phe 355 | Gly | Gly | Ser | Ser | Phe 360 | Ser | Asp | Lys | Gly | Gly 365 | Leu | Thr | Gln |
| Phe | Lys 370 | Lys | Ala | Thr | Ser | Gly 375 | Gly | Met | Val | Leu | Val 380 | Met | Ser | Leu | Trp |
| Asp 385 | Asp | Tyr | Tyr | Ala | Asn 390 | Met | Leu | Trp | Leu | Asp 395 | Ser | Thr | Tyr | Pro | Thr 400 |
| Asn | Glu | Thr | Ser | Ser 405 | Thr | Pro | Gly | Ala | Val 410 | Arg | Gly | Ser | Суз | Ser 415 | Thr |
| Ser | Ser | Gly | Val 420 | Pro | Ala | Gln | Val | Glu 425 | Ser | Gln | Ser | Pro | Asn 430 | Ala | Lys |
| Val | Thr | Phe 435 | Ser | Asn | Ile | Lys | Phe 440 | Gly | Pro | Ile | Gly | Ser 445 | Thr | Gly | Asn |

Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr Thr Arg Arg Pro Ala Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln 465 470 475 480 Ser His Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val 485 490 495 Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln 500 505 510 Cys Leu <210> 6 <211> 514 <212> PRT <213> Trichoderma reesei CBH1-N45A <400> 6 Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg 1 5 10 15 Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr 20 25 30 Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser 35 40 45 Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Ala Ser Ser 55 50 60 Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp 65 70 75 80

Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe 100 105 110

Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala

85 90 95

Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met

115 120 125

Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe 130 135 140 Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala 150 155 Leu Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro 165 170 175 Thr Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln 180 185 190 Cys Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly 195 200 205 Trp Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly 210 215 220 Ser Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu 230 235 225 Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu 245 250 255 Gly Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr 265 260 270 Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr 275 280 285 Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys 295 300 290 Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr 305 310 315 320 Tyr Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly 330 325

Ser Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu

345

340

Ala Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln 360 Phe Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp 370 375 380 Asp Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr 390 395 400 385 Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr 405 410 415 Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys 420 425 Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn 435 440 445 Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr 450 455 460 Thr Arg Arg Pro Ala Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln 470 480 465 475 Ser His Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val 485 490 495 Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln 500 505 Cys Leu <210> 7 <211> 514 <212> PRT <213> Trichoderma reesei CBH1-N270A

Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg

10 15

<400> 7

| Ala Gl | n Ser | Ala 20 | Cys | Thr | Leu | Gln | Ser 25 | Glu | Thr | His | Pro | Pro 30 | Leu | Thr |
|---------------|---------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Trp Gl | n Lys 35 | Cys | Ser | Ser | Gly | Gly 40 | Thr | Суз | Thr | Gln | Gln 45 | Thr | Gly | Ser |
| Val Va | | Asp | Ala | Asn | Trp 55 | Arg | Trp | Thr | His | Ala 60 | Thr | Asn | Ser | Ser |
| Thr As | n Cys | Tyr | Asp | Gly 70 | Asn | Thr | Trp | Ser | Ser 75 | Thr | Leu | Суз | Pro | Asp 80 |
| Asn Gl | u Thr | Cys | Ala 85 | Lys | Asn | Суз | Суз | Leu 90 | Asp | Gly | Ala | Ala | Tyr 95 | Ala |
| Ser Th | ır Tyr | Gly 100 | Val | Thr | Thr | Ser | Gly 105 | Asn | Ser | Leu | Ser | Ile 110 | Gly | Phe |
| Val Th | ır Gln 115 | Ser | Ala | Gln | Lys | Asn 120 | Val | Gly | Ala | Arg | Leu 125 | Tyr | Leu | Met |
| Ala Se | _ | Thr | Thr | Tyr | Gln 135 | Glu | Phe | Thr | Leu | Leu 140 | Gly | Asn | Glu | Phe |
| Ser Ph | ie Asp | Val | Asp | Val 150 | Ser | Gln | Leu | Pro | Cys 155 | Gly | Leu | Asn | Gly | Ala 160 |
| Leu Ty | r Phe | Val | Ser 165 | Met | Asp | Ala | Asp | Gly 170 | Gly | Val | Ser | Lys | Tyr 175 | Pro |
| Thr As | n Thr | Ala 180 | Gly | Ala | Lys | Tyr | Gly 185 | Thr | Gly | Tyr | Cys | Asp 190 | Ser | Gln |
| Cys Pr | 195 | | | | | 200 | | | | | 205 | | | |
| Trp Gl 21 | | Ser | Ser | Asn | Asn 215 | Ala | Asn | Thr | Gly | 11e 220 | Gly | Gly | His | Gly |
| Ser Cy 225 | s Cys | Ser | Glu | Met 230 | Asp | Ile | Trp | Glu | Ala 235 | Asn | Ser | Ile | Ser | Glu 240 |

Ala Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu

245 250 255

Gly Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr 260 265 270 Cys Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Ala Thr 280 285 Ser Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys 290 295 300 Leu Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr 305 310 315 320 Tyr Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly 325 330 335 Ser Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu 340 345 350 Ala Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln 355 360 365 Phe Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp 375 380 370 Asp Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr 385 390 395 400 Asn Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr 405 410 415 Ser Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys 420 425 Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn 435 440 445 Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly Thr Thr 460 455

Thr Arg Arg Pro Ala Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln

475

480

470

465

Ser His Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val 485 490 495

Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln 500 505 510

Cys Leu

<210> 8

<211> 514

<212> PRT

<213> Trichoderma reesei CBH1-N384A

<400> 8

Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg 1 5 10 15

Ala Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr 20 25 30

Trp Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser 35 40 45

Val Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser 50 60

Thr Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp 75 80

Asn Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala 85 90 95

Ser Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe 100 105 110

Val Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met 115 120 125

Ala Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe 130 135 140 Ser Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala 145 150